Vocal Tract-related Acoustic Cues Across Human and Non-Human Vocal Signals

Preprint version of:

Vocal Tract-related Acoustic Cues Across Human and Non-Human Vocal Signals. pp.1-17. In RETI SAPERI LINGUAGGI - ISSN:2279-7777 vol. 1 (1)

Marco Gamba

marco.gamba@unito.it

Vocal tract-related cues across human and nonhuman signals

Marco Gamba

Università di Torino, Dipartimento di Scienze della Vita e Biologia dei Sistemi, Via Accademia Albertina 13, 10123 Torino, Italia. Email: marco.gamba@unito.it

Abstract

I provide an overview of the studies on human speech and animal communication, showing that the formants, or vocal tract resonances, are crucial acoustic features of human and nonhuman vocalizations. The formants of nonhuman mammals possess the potential to provide conspecifics with cues to emitter's body size, species and individual identity. The findings reviewed in this paper suggest that human phonetic abilities may have evolved separately from formant perception and possibly at a later stage, as the widespread use of formants in animal communication would indicate.

Introduction

During our everyday interactions, our brain computes an incredible number of operations, including decoding and encoding the vocal sounds we hear from and exchange with others. The most prominent features of these sounds derive from a specific step in our phonation process. The vocal cues in the human voice can provide information about individuality and sex (Dellwo et al. 2007), as we experience everyday when a member of our family or a friend calls at the intercom, or when a never-before-seen call center operator phone us for a special offer. Most surprisingly similar acoustic features may play a role in the recognition of other species' calls (Molnar et al. 2006), as when we identify our family dogs and cats. These features play a role in animal communication too, providing other individuals with information about the emitter (Taylor and Reby 2010). The purpose of this paper is investigating these processes in a comparative perspective, outlining the role of formants in the vocal communication of humans and others species. Of course, what we consider the main feature of speech is its verbal information. I ask you to forget about that while reading this paper, and appreciate the nonverbal component of human speech, being ready to evaluate whether other species make a use of shared properties of the vocal signals.

The source-filter theory and human speech

Human typical sounds are shaped during two independent events taking place in our phonatory apparatus. The first of these events takes place in the larynx, when the airflow coming from the lungs elicits cyclic abduction and adduction of the vocal folds. The signal produced by this process is usually indicated as source signal (Fant 1960; Flanagan 1965) and has its salient feature in Fundamental frequency (or F_0) (Titze 1994). The recognition that larynx phonation was not enough to produce a

normal human voice is dated back to the 18th century, when von Kempelen (1791), the first experimental phonetician, was creating his speaking machine. In his machine, an ingenious mechanism made evident the role of the supralaryngeal cavities in shaping the vocal sounds. In fact, a second, independent event takes place in the vocal tract during voice production. The vocal tract, which resembles a series of tubes, acts as a resonator and filter frequencies produced at the glottis. This filtering process (Fant 1960; Flanagan 1965) enhances or leaves unchanged the intensity of the frequencies corresponding to the resonance modes of the tube and dampens the others. The description of the source-filter system presented above have been proposed by Muller (1848) and postulated by Fant (1960).

The characteristics of human speech sounds strongly depends on the resonance in the supralaryngeal airway and the filter suppresses the transfer of sound energy at certain frequencies as a consequence of its length and shape (Titze 1994).

The visual result of the filtering process, when seen on the frequency spectrum of a vowel sound, would be the presence of prominent peaks called formants. The resulting spectrum would represent the vocal signal at the end of the vocal tract (e.g. at the lips). The amplitude of each frequency component, the Fundamental frequency and its integer multiples (the harmonics), would depend on the amplitude of the specific component and from the filter function (Ladefoged 1996). Formants are usually numbered subsequently according to their increasing frequency. They represent the centre frequencies of the supralaryngeal vocal tract filter (Lieberman and Blumstein 1988). Formants measured in human speech sounds are usually three or four (F_1, F_2, F_3, F_4) (Boersma 2014).

Voice perception in humans follows very complex neural mechanisms (Miller 1990) that I will not investigate here. But we know that the pitch (a perceptual

representation of the Fundamental frequency) of a speaker's voice is important to perception of various speaker characteristics (e.g. gender, Gelfer and Bennett 2013) and that that Fundamental frequency differences are crucial to segregate acoustic/phonetic components during perception of a speech signal (Darwin and Gardner 1986; Gardner et al. 1989).

Previous studies have shown that male and female voices differ in their Fundamental frequency (Aronovich 1976; Banse and Scherer 1996). In the comparative analysis on European languages by Traunmüller and Eriksson (1994), males showed lower Fundamental frequency (119 Hz) than females (207 Hz). Fundamental frequency differences among men and women can be explained in relation to differences in larynx size and vocal fold length. Kahane (1978) and Hirano and colleagues (1983) showed that the human larynx and the length of the vocal folds increased during development and stops at around twenty years old, when male larynx is approximately 20% bigger than women's one. At that age, the membranous length of the vocal folds is approximately 10 mm in men and 16 mm in women. Titze (1989) found that this factor (1.6) accounts entirely for differences in mean Fundamental frequency of men and females sounds.

Formants are not only the key acoustic parameters defining differences between phonemes in speech, showing that vocal gestures are significantly impacting vocal tract resonance, but also provide listeners with various cues about speaker's characteristics. A simple empirical experiment can be done when speaking in front of a mirror. We could then appreciate how the vocal tract length can be varied through adjustments of larynx position, mouth opening and lip protrusion.

Formant measurements show clear sex and age differences in the resonance of men, women and children (Chiba and Kajiyama 1941; Potter and Steinberg 1950; Peterson

and Barney 1952). These differences are consistent for different vowel categories. They have generally been interpreted as a consequence of different vocal tract size. Changes in the vocal tract length during ontogeny (Fitch and Giedd 1999; Vorperian and Kent 2007) have dramatic effects on formant frequencies. Formants, and especially the spacing between formant, usually decrease as the vocal tract lengthens. Values usually considered as an average are 6 to 8 cm for neonates, 15 cm for adult females and 18 cm for adult males. It can be of interest to notice whether no final evidence of sexual dimorphism in Fundamental frequency until puberty (Hirano et al. 1981), the formant pattern may be useful to provide sex specific information in pre-pubertal subjects (Cartei and Reby 2013).

Formants are known to inform about individuality because many possible adjustments occurring in the phonatory apparatus, where the physical size of the organs is only one source of variation, but important (see Dellwo et al. 2007). The individual-specific length of the vocal tract and the size of the cavities affect acoustic resonance. Additionally, I have already mentioned the effects of lips and larynx movements. Listening to someone speaking may allow us inferring about the speaker's sex (Hillebrand and Clarke 2009), regional origin (Clopper and Pisoni 2004a; 2004b; Jacewicz et al. 2011) and, to a limited extent, age (Schötz and Svantesson 2005), body size (Rendall et al. 2007) and emotional state (Banse and Scherer 1996; Goudbeek et al. 2009; Khulage and Pathak 2012).

The role of formants in animal communication

Despite the integration between linguistic, phonetics studies and animal behaviour and communication has made relevant steps in recent years, the parallel view of typical acoustic cues of human speech in relation to those found in animal

communication is still at the very beginning. Briefly, taking the same perspective used above, let me consider that whether morphology of the mammalian larynx has shown moderate variability (Negus 1949; Harrison 1995), its basic layout is very conservative among most terrestrial mammals (Fitch 2000a). Keeping my focus on formants, I will not get into the details of studies concerning larynx-related vocal cues in animal communication. Let me just summarize that whether it has been demonstrated that the lowest deliverable F_0 of an individual is correlated with the length and mass of its vocal folds, the correlation between the vocal fold length and the caller's size is not robust (Fitch and Hauser 1995; 2002; Reby and McComb 2003; Rendall et al. 2005; Sanvito et al. 2007).

The study of formants in non-human species has received discontinued attention from researchers. Nonetheless, the evidence that formants may play a role in animal communication and that they possess the potential to convey information in animal systems is solid and constantly growing. In the last century, the study of vocal tract resonance in non-human primates started as an attempt to demonstrate that primates could not produce human speech sounds. Once the pioneering works of Lieberman and colleagues (Lieberman 1968; 1969; Lieberman et al. 1969; 1972) showed that the anatomy of the non-human primate vocal tract prevents the production of human-like sounds, this field of investigation has seen dramatic loss of interest. Later on, the work of Andrew (1976) and then Hauser (Hauser et al. 1993; Hauser 1996) brought back attention to the meaning of formants in primate intra-specific communication. The studies conducted by Marc Hauser (Hauser 1991; 1992; 1996; Hauser and Fowler 1992; Hauser et al. 1993) showed that monkeys and apes possess the ability to modify the shape and the length of their vocal tract. The point is not whether other species can produce the wide variety of sounds typically used in human speech, but to

understand how these species may change the spectral properties of their calls and how these changes determine crucial features of their vocal communication. An influential paper in primate formant studies is the work of Fitch on the correlation between body size and vocal tract length in macaques (1997). In this paper, Fitch demonstrates the correlation between the distance between formants, vocal tract length and body size. Formant position and spacing provided an accurate means of predicting a vocalizer's body length and mass. Publications by Fitch and Hauser (Fitch and Hauser 1995; Fitch 1997), together with the papers by Owren, Rendall and colleagues (Owren 1990; Owren et al. 1993; Owren and Linker 1995; Owren et al. 1997; Rendall et al. 1996; Rendall et al. 1998; Rendall et al. 1999) were the first moves towards the quantitative investigation of vocal tract-related acoustic cues in animal communication in the last two decades.

These studies, among others, cleared out the sometimes apparent sometimes concealed idea that variable formant patterns would be allowed only for organisms showing the larynx in a "lowered" (or "descended") position. Thus, only humans would be able to produce proper formants. Actually many studies showed that, while the larynx in many species is inserted into the nasal passages when in resting position, it could be temporarily lowered during phonation (Fitch 2000b). This is the case of lemurs, which can perform (similarly to human neonates) a nasal/laryngeal seal. They lower the larynx to produce oral calls and slightly move it to produce nasal calls (Gamba and Giacoma 2006; Gamba et al. 2012). This result was then complemented by studies showing that some species possess a permanently lowered larynx (e.g. deers or Mongolian gazelles (Fitch and Reby 2001; Weissengruber et al. 2002).

The vocal tract in most terrestrial mammals is confined in the head, thus its length is strongly correlated with skull length, and thus with body size (Fitch 1997; 2000c).

Because of this, the formant frequencies are feasibly providing conspecifics with honest indications of the emitter's body size (Fitch and Hauser 1995). Of course, this has important implications for sexual selection and mate competition and, in fact, it has been demonstrated for many species (Fitch 1997; Harris et al. 2006; Sanvito et al. 2007; Vannoni and McElligott 2008; Charlton et al. 2009), in which body size and skull size are correlated. The link between formants and body size could have favoured the evolution of morphological adaptations directed to elongate the vocal tract and invalidate the honesty of the signal (Fitch and Hauser 1995). The ability of increase the length of the vocal tract would result in broadcasting the information of a larger emitter. Context-specific systematic use of the longer available tract (e.g. the nasal tract) would maximize the possible cues to body size provided to listeners (Gamba, unpublished data). But, as mentioned above, there are also other strategies for vocal tract elongation, as lowering the larynx during vocal production (e.g. deer, Fitch and Reby 2001; Reby and McComb 2003; koalas, Charlton et al. 2011; dogs, Riede and Fitch 1999).

The absence of a proper descended larynx has not limited other mammalian species to use articulators to modify the vocal tract resonance. Whereas the role of tongue movements is still unclear (Riede et al. 2005; Lieberman 2006; Fitch and Zuberbühler 2013), mouth opening and lip position can be both used to alter the sound (Gamba and Giacoma 2006; Favaro et al. 2008; Gamba et al. 2011). I was strongly impressed to realize that indris tune their vocal tract in a way that favours the landing of the fundamental frequency of their calls in the frequency window of the first formant (Favaro et al. 2008; Gamba et al. 2011). Although further studies are needed to clarify the structure of the vocal tract in these species, I am confident that the most important determinant of the indris formant patterns is mouth opening and that this strategy is

used in order to increase the broadcast area of the signalling in its dense tropical rain forest environment. This is strikingly similar to what has been suggested for female human singers, who share with indris the remarkably high pitch of their songs (Sorrentino et al. 2013), producing powerful high tones to be heard over the orchestra (Sundberg 1987; Joliveau et al. 2004).

Going back to indexical cues in primate calls, I want to mention some studies focusing on the potential for species- and individual-recognition. The species-specific potential of formant patterns in nasal vocalizations of *Eulemur* species was investigated in Gamba, Friard and Giacoma (2012) by means of acoustic analysis and computational models of the vocal tract. This study allowed understanding that the shape and length of the lemur nasal tract are both influencing the acoustic output and that the formants resulting from resonance on the nasal airways have a potential for species recognition. The combined use of modeling and quantitative acoustic analyses demonstrated that laryngeal morphology and the nasal airways determine differences in formant position and formant spacing in lemurs. This study also showed that prominent spectral features of lemur calls are indeed the result of vocal tract filtering. Another paper on lemurs (*Eulemur rubriventer*, Gamba et al. 2012) showed that formants measured from nasal grunt of different individuals have distinctive patterns and thus are potentially providing conspecifics with individual cues. These results are in agreement with the data presented by Rendall and colleagues (1996), who suggested that formant information in macaque provides conspecifics with a relatively stable indicator of individual identity.

There is more ambiguity when we analyse sex- and age-related cues. The spacing between the first and the second formant decreased with age in baboon barks (Fischer

et al. 2002; Fischer et al. 2004) but has seen controversial result in Pfefferle and Fischer (2006) where they failed to find these correlation in adult females.

Although it is likely to be present in many species, we still lack evidence about sexual dimorphism in formant patterns. The relationship between sex and formants was rarely investigated (Ey et al. 2007). Potential sex information has been found in the *clicks* of crowned lemurs (*Eulemur coronatus*), where Discriminant analysis identified statistically significant differences in the third and the fourth formant (Gamba and Giacoma 2007). Available data show very little sexual dimorphism in the body size of *Eulemur coronatus* (Leigh and Terranova 1998) and sex dimorphism has been found in one call type only. The above-mentioned studies fits well with those showings that formant-based semantic communication is present in non-human primates (Owren and Bernacki 1998; Riede and Zuberbuhler 2003a, b; Rendall et al. 2005).

I want to highlight that the importance of vocal tract filtering is not limited to primates. Songbirds produce acoustically complex utterances that exhibit rapid and dramatic variations in frequency and duration. Several findings indicated that vocal tract resonance have a role in bird communication. Vocal tract shape and length modify the sound produced at the syrinx (Gaunt and Nowicki 1997). Beak and tongue movements, and neck elongation were known to emphasize the time-frequency patterns of the sounds produced at the syrinx but their significance in changing the configuration of the vocal tract and determine the acoustic structure of birdcalls has been widely recognized only in recent years. Tongue movements observed in parrot vocalizations modulate formant characteristics independently from the vocal source (Nottebohm 1976; Patterson and Pepperberg 1998) showing a striking similarity with a speech-like formant system (Beckers et al., 2004). The birdsong vocal system can actually be much more complex than previously though. Riede, Suthers, Fletcher, and

Blevins (2005) studied the impacts of beak gape, the oropharyngeal–esophageal filter and the tracheal filter, have been studied experimentally using cineradiography and computational modelling. As reported above for the indris, thanks to a complex motor coordination can adjust their vocal tract using strategies similar to those shown by human sopranos.

Birds, primates and other mammals may actually use their supra-laryngeal tracts to modify their vocalizations and may change the configuration of their airways to shape the formant patterns of their calls to a significant extent.

A comparative perspective

I showed that spectral prominences in the vocalizations of a growing number of nonhuman species are formants. These parameters have the potential to provide conspecific listeners with honest information on some emitter's characteristics similarly to what we are used to in human vocal communication. But the potential of providing this information does not necessarily correspond with the receiver's ability to perceive the signal. Receiver preadaptation plays an important role in the evolution of signals, including vocalizations. Thus, we should ask whether listeners could perceive formants. Three studies can provide interesting data to this point. First, spontaneous formant perception has been demonstrated in macaques (Fitch and Fritz 2006). Second, the ability of baboons (*Papio anubis*) to discriminate changes in formant structure is roughly comparable to the human ability (Hienz et al. 2004). Third, the use of playback experiments allowed demonstrating that female red deer are sensitive to shifts in formant frequencies (Charlton et al. 2007).

These studies indicated that formants are salient for nonhuman mammals and that they have a strong biological significance. Thus, other species can perceive formants

and can discriminate formant changes similarly to that observed in humans. I believe that the data presented above indirectly support the hypothesis that formants can be perceived spontaneously in Strepsirrhini (Gamba et al. 2012), as demonstrated in macaques. Evaluating these results in the light of the perceptual abilities of lemurs, Gamba, Friard and Giacoma (2012) found that the second, third and fourth formants are comprised in the optimal sensitivity of *Eulemur fulvus* and *Eulemur macaco* (Mitchell et al. 1970). These formants were also more stable across species and appeared to have a stronger correlation with vocal tract length when compared to the first formant. Gamba, Friard and Giacoma (2012) also showed that between-species formant variation is around 15% for the second and third formants, and around 8% between conspecifics. When these results are compared with those reported by Kewley-Port and Watson (1994) about human voice the reader could notice that formant changes can be discriminated by lemurs in a way that may be not fine-scaled as the one of humans, but it is roughly similar.

Consequently, we can imagine that, at least in several other mammal species, individuals may regulate their social interactions and adjust their behavioural responses according to vocal signals in which formants are meaningful features.

Broadening our view, we can infer about the evolution of formant communication. The evidence concerning formant perception and production in nonhuman mammals suggests that it should have evolved independently from larynx position. Species without permanently descended larynges use formants as vocal cues. As argued by Fitch and Reby (2001), larynx position in the vocal tract may not have evolved for phonetic functions and a phonetic-driven analysis of vocal tract morphology based solely on this character must be undertaken in a cagily. The widespread ability of elongating the vocal tract (Fitch 1999) may have been evolved under the pressure of

body size exaggeration and not to increase the articulatory potential and widen the vocal repertoire (Fitch 1997; Fitch and Giedd 1999).

The human two-tube vocal tract configuration, possibly not uniquely human to a certain extent (Nishimura et al. 2003; Nishimura et al. 2008; Fitch and Zuberbühler 2013), surely enables a wider range of resonance patterns and complex articulation when compared to any other mammalian vocal system. But there are many other adaptations. Some species show a permanently descended larynx, others present species-typical modifications of the phonatory apparatus, like an air sac (e.g. siamangs) or a short stump of a nose (e.g. snub-nosed monkeys). Still others, like some birds, have the capacity to learn vocalizations. Future works will shed light on the selective pressures that might have shaped these processes. But the use of formant communication appears as an ancient aspect of vocal communication, which possibly pre-dated the two tube vocal tract and anticipated the uniquely human phonetic abilities.

Reference

Andrew R.J. (1976), *Use of formants in the grunts of baboons and other nonhuman primates*, in «Annals of the New York Academy of Sciences», 280, 673-679.

Aronovich C.D. (1976), *The voice of personality: stereotyped judgments and their relation to voice quality and sex of speaker*, in «Journal of Social Psychology», 99, 207–220.

Banse R., Scherer K.R. (1996), *Acoustic profiles in vocal emotion expression*, in «Journal of Personality and Social Psychology», 70, 614–636.

Beckers G.J.L., Nelson B.S., Suthers R.A. (2004), *Vocal tract filtering by lingual articulation in a parrot*, in «Current Biology», 14, 1592-1597.

Boersma P. (2014), *Acoustic analysis*, in Podesva R., Sharma D. (eds.), *Research methods in linguistics*, Cambridge, Cambridge University Press.

Cartei V., Reby D. (2013), *Effect of formant frequency spacing on perceived gender in pre-pubertal children's voices*, in «PLoS ONE», 8(12), e81022.

Charlton B.D., Ellis W.A.H., McKinnon A.J., Cowin G.J., Brumm J., Nilsson K., Fitch W.T. (2011), *Cues to body size in the formant spacing of male koala (Phascolarctos cinereus) bellows: honesty in an exaggerated trait*, in «Journal of Experimental Biology»214, 3414-3422.

Charlton B.D., Reby D., McComb K. (2007), *Female perception of size-related formant shifts in red deer, Cervus elaphus*, in «Animal Behaviour», 74 (4), 707-714.

Charlton B.D., Zhihe Z., Snyder R.J. (2009), *The information content of giant panda, Ailuropoda melanoleuca, bleats: acoustic cues to sex, age and size*, in «Animal Behaviour», 78, 893-898.

Chiba T., Kajiyama M. (1941) *The vowel: its nature and structure*, Tokyo, Tokyo Publishing Company.

Clopper C.G., Pisoni D.B: (2004a), *Effects of talker variability on perceptual learning of dialects*, in «Language and Speech» 47(3), 207-238.

Clopper C.G., Pisoni D.B. (2004b), *Some acoustic cues for the perceptual categorization of American English regional dialects*, in «Journal of Phonetics», 32(1), 111-140.

Darwin C.J., Gardner R.B. (1986), *Mistuning a harmonic of a vowel: grouping and phase effects on vowel quality*, in «Journal of the Acoustical Society of America», 79, 838-845.

Dellwo V., Huckvale M., Ashby M. (2007), *How is individuality expressed in voice? An introduction to speech production & description for speaker classification*, in C. Müller (ed.), *Speaker classification I*, Berlin, Springer Verlag, 1-20.

Ey E., Pfefferle D., Fischer J. (2007), *Do age- and sex-related variations reliably reflect body size in non-human primate vocalizations? A review*, in «Primates», 48, 253-267.

Fant G. (1960), *Acoustic theory of speech production*, The Hague, Mouton.

Favaro L., Gamba M., Sorrentino V., Torti V., Giacoma C. (2008), *Singers in the forest: acoustic structure of indri's loud calls and first evidence of vocal tract tuning in a prosimian primate*, in «Atti del 35° Convegno Nazionale dell'Associazione Italiana di Acustica», Milano, Associazione Italiana di Acustica.

Fischer J., Hammerschmidt K., Cheney D.L., Seyfarth R.M. (2002), *Acoustic features of male baboon loud calls: influences of context, age, and individuality*, in «The Journal of the Acoustical Society of America», 111, 1465-1474.

Fischer J., Kitchen D.M., Seyfarth R.M., Cheney D.L. (2004), *Baboon loud calls advertise male quality: acoustic features and their relation to rank, age, and exhaustion*, in «Behavioral Ecology and Sociobiology», 56(2), 140-148.

Fitch W. T. (2000c), *Skull dimensions in relation to body size in nonhuman mammals: the causal bases for acoustic allometry*, in «Zoology», 103, 40-58.

Fitch W.T. (1997), *Vocal tract length and formant frequency dispersion correlate with body size in rhesus macaques*, in «Journal of the Acoustical Society of America», 102, 1213-1222.

Fitch W.T. (1999), Acoustic exaggeration of size in birds by tracheal elongation: comparative and theoretical analyses, in «Journal of Zoology (London)», 248, 31-49.

Fitch W.T. (2000a), *The evolution of speech: a comparative review*, in «Trends in Cognitive Sciences», 4, 258-267.

Fitch W.T. (2000b), *The phonetic potential of nonhuman vocal tracts: comparative cineradiographic observations of vocalizing animals*, in «Phonetica», 57 (2-4), 205-218.

Fitch W.T., Fritz J.B. (2006), *Rhesus macaques spontaneously perceive formants in conspecific vocalizations*, «The Journal of the Acoustical Society of America», 120, 2132-2141.

Fitch W.T., Giedd J. (1999), *Morphology and development of the human vocal tract: A study using magnetic resonance imaging*, in «Journal of the Acoustical Society of America», 106, 1511–1522.

Fitch W.T., Hauser M.D. (1995), *Vocal production in nonhuman primates: acoustics, physiology and functional constraints on honest advertisement*, in «American Journal of Primatology», 7, 191-219.

Fitch W.T., Hauser M.D. (2002), *Unpacking "honesty": generating and extracting information from acoustic signals*, in A. Megela-Simmons, A. Popper (eds.), *Animal Communication*, Berlin, Springer-Verlag, 65-137.

Fitch W.T., Reby D. (2001), *The descended larynx is not uniquely human*, in «Proceedings of the Royal Society Biological Sciences» 268, 1669–1675.

Fitch W.T., Zuberbühler K. (2013), *Primate precursors to human language: beyond discontinuity*, in E. Altenmüller, S. Schmidt, E. Zimmermann (eds.), *The evolution of emotional communication: from sounds in nonhuman mammals to speech and music in man*, 26-48.

Flanagan J.L. (1965), *Speech analysis synthesis and perception*, New York, Springer-Verlag.

Gamba M., Colombo C., Giacoma C. (2012), *Acoustic cues to caller identity in lemurs: a case study*, in «Journal of Ethology», 30(1), 191-196.

Gamba M., Favaro L., Torti V., Sorrentino V., Giacoma C. (2011), *Vocal tract flexibility and variation in the vocal output in wild indris*, in «Bioacoustics», 20(3), 251-265.

Gamba M., Friard O., Giacoma C. (2012), *Vocal tract morphology determines species-specific features in vocal signals of lemurs (Eulemur)*, in «International Journal of Primatology», 33, 1453-1466.

Gamba M., Giacoma C. (2006), *Vocal tract modeling in a prosimian primate: the black and white ruffed lemur*, in «Acta Acustica United with Acustica», 92, 749-755.

Gamba M., Giacoma C. (2007), *Quantitative acoustic analysis of the vocal repertoire of the crowned lemur*, in «Ethology, Ecology and Evolution», 19(4), 323-343.

Gardner R.B., Gaskill S.A., Darwin C.J. (1989), *Perceptual grouping of formants with static and dynamic differences in fundamental frequency*, in «Journal of the Acoustical Society of America», 85, 1329-1337.

Gaunt A.S., Nowicki S. (1998), *Sound production in birds: acoustics and physiology revisited*, in S.L. Hopp, M.J. Owren, C.S. Evans (eds.), *Animal acoustic communication*, Berlin, Springer, 291-321.

Gelfer M.P., Bennett Q. (2013), *Speaking fundamental frequency and vowel formant frequencies: effects on perception of gender*, in «Journal of Voice», 27, 556-566.

Goudbeek M., Goldman J.P., Scherer K.R. (2009), *Emotion dimensions and formant position*", in «Interspeech-2009», 1575-1578.

Harris T.R., Fitch W.T., Goldstein L.M., Fashing P.J. (2006), *Black and white colobus monkey (Colobus guereza) roars as a source of both honest and exaggerated information about body mass*, in «Ethology», 112, 911-920.

Harrison D.F.N. (1995), *The Anatomy and Physiology of the Mammalian Larynx*, Cambridge, Cambridge University Press.

Hauser M.D. (1991), *Sources of acoustic variation in rhesus macaque vocalizations*, in «Ethology», 89: 29-46.

Hauser M.D. (1996), *The evolution of communication*, Cambridge, MIT Press.

Hauser M.D., Evans C.S., Marler P. (1993), *The Role of Articulation in the Production of Rhesus Monkey (Macaca mulatta) Vocalizations*, in «Animal Behaviour», 45, 423-433.

Hauser M.D., Fowler C. (1992), *Declination in fundamental frequency is not unique to human speech: evidence from nonhuman primates*, in «Journal of the Acoustical Society of America», 91: 363- 369.

Hauser, M.D. (1992), *Articulatory and social factors influence the acoustic structure of rhesus monkey vocalizations: A learned mode of production?*, in «Journal of the Acoustical Society of America», 91, 2175- 2179.

Hienz R.D., Jones A.M., Weerts E.M. (2004), *The discrimination of baboon grunt calls and human vowel sounds by baboons*, in «Journal of the Acoustical Society of America», 116, 1692–1697.

Hillenbrand JM, Clark MJ (2009), *The role of F0 and formant frequencies in distinguishing the voices of men and women*, in «Attention, Perception, & Psychophysics», 71, 1150–1166.

Hirano M., Kurita S., Nakashima T. (1981), *The structure of the vocal folds*, in K. Stevens, M. Hirano (eds.), *Vocal Fold Physiology*. Tokyo, University of Tokyo Press, 33–43.

Hirano M., Kurita S., Nakashima T. (1983), *Growth, development and aging of human vocal folds*, in D.M. Bless, J.H. Abbs (eds.), *Vocal Fold Physiology*, San Diego, College-Hill Press, 22-43.

Jacewicz E., Fox R.A., Salmons J. (2011), *Regional dialect variation in the vowel systems of typically developing children*, in «Journal of Speech, Language and Hearing Research», 54 (2), 448-470.

Joliveau E., Smith J., Wolfe J. (2004) *Tuning of vocal tract resonance by sopranos*, in «Nature», 427, 116.

Kahane J. (1978), *A morphological study of the human prepubertal and pubertal larynx*, in « American Journal of Anatomy», 151, 11-20.

Kewley-Port D., Watson C.S. (1994), *Formant-frequency discrimination for isolated English vowels*, in «Journal of the Acoustical Society of America», 95, 485–496.

Khulage A.A., Pathak B.V. (2012), *Analysis of speech under stress using linear techniques and non-Linear techniques for emotion recognition system*, in «The Computing Research Repository», 1-10.

Ladefoged P. (1962), *Elements of acoustic phonetics*, Chicago, University of Chicago Press.

Leigh S.R., Terranova C.J. (1998), *Comparative perspectives on bimaturism, ontogeny, and dimorphism in lemurid primates*, in «International Journal of Primatology», 19(4), 723-749.

Lieberman P. (1968), *Primate vocalizations and human linguistic ability*, in «Journal of the Acoustical Society of America», 44, 1157-1184.

Lieberman P. (1969), *On the acoustic analysis of primate vocalizations*, in «Behavior Research Methods and Instruments», 1, 169-174.

Lieberman P. (2006), *Limits on tongue deformation. Diana monkey formants and the impossible vocal tract shapes proposed by Riede et al. (2005)*, in «Journal of Human Evolution», 50, 219-221.

Lieberman P., Blumstein S. (1988), *Speech physiology, speech perception, and acoustic phonetics*, Cambridge, Cambridge University Press.

Lieberman P., Harris K.S., Wolff P., Russell L.H. (1972), *Newborn infant cry and nonhuman-primate vocalizations*, in «Journal of Speech and Hearing Research», 14, 718-727.

Lieberman P., Klatt D.H., Wilson W.H. (1969), *Vocal tract limitations on the vowel repertoires of rhesus monkey and other nonhuman primates*, in «Science», 164, 1185-1187.

Miller J.L. (1990), *Speech perception*, in D.N. Osherson, H. Lasnik (eds.), *An invitation to cognitive science*, Cambridge, MIT Press, 69-93.

Mitchell C., Gillette R., Vernon J., Herman P. (1970), *Pure-tone auditory behavioral thresholds in three species of lemurs*, in «Journal of the Acoustical Society of America», 48, 531–535.

Molnár C., Pongrácz P., Dóka A., Miklósi A. (2006), *Can humans discriminate between dogs on the base of the acoustic parameters of barks?*, in «Behavioural processes», 73(1), 76-83.

Muller J. (1848), *The physiology of the senses, voice and muscular motion with the mental faculties*, London, Walton and Maberly.

Negus V. E. (1949), *The Comparative Anatomy and Physiology of the Larynx*, London, William Heinemann Medical Books.

Nishimura T., Mikami A., Suzuki J., Matsuzawa T. (2003), *Descent of the larynx in chimpanzee infants*, in «Proceedings of the National Academy of Sciences of the United States of America», 100(12), 6930-6933.

Nishimura T., Oishi T., Suzuki J., Matsuda K., Takahashi T. (2008), *Development of the supralaryngeal vocal tract in Japanese macaques: implications for the evolution of the descent of the larynx*, in «American Journal of Physical Anthropology», 135(2), 182-194.

Nottebohm F. (1976), *Phonation in the orange-winged amazon parrot, Amazona amazonica*, in «Journal of Comparative Physiology», 108, 157–170.

Owren M.J. (1990), *Acoustic classification of alarm calls by vervet monkeys (Cercopithecus aethiops) and humans (Homo sapiens): I. Natural calls*, in «Journal of Comparative Psychology», 104(1), 20-28.

Owren M.J., Bernacki R.H. (1998), *Applying linear predictive coding (LPC) to frequency-spectrum analysis of animal acoustic signals*, in S.L. Hopp, M.J. Owren, C.S. Evans (eds.), *Animal acoustic communication*, Berlin Heidelberg, Springer Verlag, 129-162.

Owren M.J., Dieter J.A., Seyfarth R.M., Cheney D.L. (1993), *Vocalizations of rhesus (Macaca mulatta) and Japanese (M. Fuscata) macaques cross-fostered between species show evidence of only limited modification*, in «Developmental Psychobiology», 26(7), 389-406.

Owren M.J., Linker C.D. (1995), *Some analysis methods that may be useful to acoustic primatologists*, in U. Jürgens, J. Newman, E. Zimmermann (eds.), *Current topics in primate vocal communication*, 1-27.

Owren M.J., Seyfarth R.M., Cheney D.L. (1997), *The acoustic features of vowel-like grunt calls in chacma baboons (Papio cynocephalus ursinus): implications for production processes and functions*, in «Journal of the Acoustical Society of America», 101, 2951-2963.

Patterson D.K., Pepperberg I.M. (1998), *Acoustic and articulatory correlates of stop consonants in a parrot and a human subject*, in «Journal of the Acoustical Society of America», 103, 2197–2215.

Peterson G., Barney H.L. (1952), *Control methods used in a study of the vowels*, in «Journal of the Acoustical Society of America», 24, 175–184.

Pfefferle D., Fischer J. (2006), *Sounds and size: identification of acoustic variables that reflect body size in hamadryas baboons, Papio hamadryas*, in «Animal Behaviour», 72, 43-51.

Potter R.K., Steinberg J.C. (1950), *Toward the specification of speech*, in «Journal of the Acoustical Society of America», 22, 807–820.

Reby D., McComb K. (2003), *Anatomical constraints generate honesty: acoustic cues to age and weight in the roars of red deer stags*, in «Animal Behaviour», 65(3), 519-530.

Rendall D., Kollias S., Ney C., Lloyd P. (2005), *Pitch (F0) and formant profiles of human vowels and vowel-like baboon grunts: The role of vocalizer body size and voice-acoustic allometry*, in «Journal of the Acoustical Society of America», 117, 944-955.

Rendall D., Owren M.J., Rodman P.S. (1998), *The role of vocal tract filtering in identity cueing in rhesus monkey (Macaca mulatta) vocalizations*, in «Journal of the Acoustical Society of America», 103, 602-614.

Rendall D., Rodman P.S., Emond R.E. (1996), *Vocal recognition of individuals and kin in free-ranging rhesus monkeys*, in «Animal Behaviour», 51, 1007-1015.

Rendall D., Seyfarth R.M., Cheney D.L., Owren M.J. (1999), *The meaning and function of grunt variants in baboons*, in «Animal Behaviour», 57, 583-592.

Rendall D., Vokey J.R., Nemeth C. (2007), *Lifting the curtain on the Wizard of Oz: biased voice-based impressions of speaker size*, in «Journal of Experimental Psychology Human Perception and Performance », 33, 1208–1219.

Riede T., Bronson E., Hatzikirou H., Zuberbühler K. (2005), *Vocal production mechanisms in a non-human primate: morphological data and a model*, in «Journal of Human Evolution», 48(1), 85-96.

Riede T., Fitch W.T. (1999), *Vocal tract length and acoustics of vocalization in the domestic dog (Canis familiaris)*, in «Journal of Experimental Biology», 202, 2859-2867.

Riede T., Suthers R.A., Fletcher N.H., Blevins W.E. (2006), *Songbirds tune their vocal tract to the fundamental frequency of their song*, in «Proceedings of the National Academy of Sciences», 103(14), 5543-5548.

Riede T., Zuberbühler K. (2003a), *Pulse register phonation in Diana monkey alarm calls,* in «Journal of the Acoustical Society of America», 113, 2919-2926.

Riede T., Zuberbühler K. (2003b), *The relationship between acoustic structure and semantic information in Diana monkey alarm vocalization*, in « Journal of the Acoustical Society of America» 114, 1132-1142.

Sanvito S., Galimberti F., Miller E.H. (2007), *Vocal signalling in male southern elephant seals is honest but imprecise*, in «Animal Behaviour», 73, 287-299.

Schötz S., Svantesson J.-O. (2005), *Automatic prediction of speaker age using CART*, Lund, Lund University.

Sorrentino V., Gamba M., Giacoma C. (2013), *A quantitative description of the vocal types emitted in the indri's song*, in J.C. Masters, M. Gamba, F. Genin (eds.), *Leaping Ahead: Advances in Prosimian Biology*, New York, Springer Science + Business Media.

Sundberg J. (1987), *The science of the singing voice*, Dekalb, Northern Illinois University Press.

Taylor A.M., Reby D. (2010), *The contribution of source–filter theory to mammal vocal communication research*, in «Journal of Zoology», 280(3), 221-236.

Titze I.R. (2000), *Principles of voice production*. Englewood Cliffs, Prentice Hall.

Titze, I.R. (1989). *Physiologic and acoustic differences between male and female voices*, in «Journal of the Acoustical Society of America», 85(4), 1699-1707.

Traunmüller H., Eriksson A. (1994), *The frequency range of the voice fundamental in the speech of male and female adults*, Stockholm, University of Stockholm.

Vannoni E., McElligott A.G. (2008), *Low frequency groans indicate larger and more dominant fallow deer (Dama dama) males*, in «PLoS ONE 3», e3113.

Von Kempelen W. (1791), *Mechanismus Der Menschlichen Sprache Nebst Beschreibung Seiner Sprechenden Maschine*, Wien, J.B. Degen.

Vorperian H.K., Kent R.D. (2007), *Vowel Acoustic Space Development in Children: A Synthesis of Acoustic and Anatomic Data*, in «Journal Speech Language and Hearing Research», 50(6), 1510–1545.

Weissengruber G.E., Forstenpointner G., Peters G., Kübber-Heiss A., Fitch W.T. (2002), *Hyoid apparatus and pharynx in the lion (Panthera leo), jaguar (Panthera onca), tiger (Panthera tigris), cheetah (Acinonyx jubatus), and domestic cat (Felis silvestris f. catus)*, in «Journal of Anatomy», 201, 195–209.